

*Electronic Letters on Computer Vision and Image Analysis 14(3):59-60, 2015*

# Segmentation and indexation of complex objects in comic book images

Christophe Rigaud<sup>\*+</sup>

<sup>\*</sup> *L3i Laboratory, University of La Rochelle, Avenue Michel Crépeau, 17042 La Rochelle, France*

<sup>+</sup> *Computer Vision Center, Universitat Autònoma de Barcelona, Edificio O - Campus UAB, 08193 Bellaterra, Spain*

*Advisors: Jean-Christophe Burie<sup>\*</sup>, Dimosthenis Karatzas<sup>+</sup> and Jean-Marc Ogier<sup>\*</sup>*

*Date and location of PhD thesis defense: 11<sup>th</sup> December 2014 at University of La Rochelle in France*

Received 13th November 2015; accepted 23rd December 2015

---

## Abstract

Born in the 19th century, comics are a visual medium used to express ideas via images, often combined with text or visual information. It is an art form that uses images deployed in sequence for graphic storytelling (sequential art), spread worldwide initially using newspapers, books and magazines. Nowadays, the development of the new technologies and the World Wide Web is giving birth to a new form of paperless comics that takes advantage of the virtual world freedom. However, traditional comics still represent an important cultural heritage in many countries. They have not yet received the same level of attention as music, cinema or literature about their adaptation to the digital format. Using information technologies with digitized comic books would facilitate the exploration of digital libraries, accelerate their translation, allow augmented reading, speech playback for the visually impaired *etc.*

Heritage museums such as the CIBDI (French acronym for International City of Comic books and Images), the Kyoto International Manga Museum and The Digital Comic Museum have already digitized several thousands of comic albums that some are now in the public domain. Despite the growing market place of digital comics, few research has been carried out to take advantage of the added value provided by these new media. A particularity of documents is their dependence on the type of document that often requires specific processing. The challenge of document analysis systems is to propose generic solutions for specific problems. The design process of comics is so specific that their automated analysis may be seen as a niche research field within document analysis, at the intersection of complex background, semi-structured and mixed content documents. Being at the intersection of several fields combines their difficulties. In this thesis, we review, highlight and illustrate the challenges related to comic book image analysis in order to provide a good overview about the last research progress in this field and the current issues. In order to cover the widest possible scope of study, we propose three different approaches for comic book image analysis. The three approaches aim to provide an automatic description of the image content. Different levels of description are discussed, from spacial positions

---

Correspondence to: [<christophe.rigaud@univ-lr.fr>](mailto:christophe.rigaud@univ-lr.fr)

Recommended for acceptance by Jorge Bernal

DOI <http://dx.doi.org/10.5565/rev/elcvia.833>

ELCVIA ISSN:1577-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

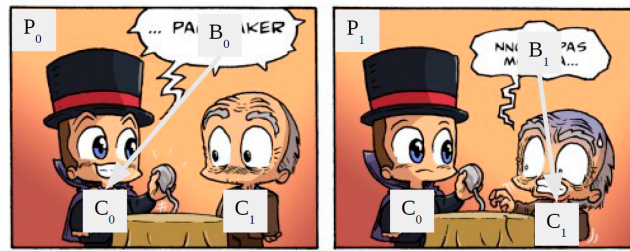


Figure 1: Low level description (positions of panels, balloons and comic characters) combined with semantic description. The left panel  $P_0$  represents a comics character labeled as  $C_0$  saying the content of balloon  $B_0$  to another character labeled as  $C_1$ . In the right panel  $P_1$ , the character  $C_1$  is answering the content of  $B_1$  to  $C_0$ .

(low level) to semantic information (high level), see Figure 1.

The first approach describes the image in an intuitive way, from simple to complex elements using previously extracted elements to guide further processing. Simple elements such as panel, text and balloon regions are extracted first, followed by balloon tails and comic character positions from the direction indicated by the tails. The second approach addresses independent information extraction to recover the main drawback of the first approach: error propagation. This second method is composed by several specific extractors for each type of content, independent from each other. Those extractors can be used in parallel, without needing previous information which cancels the error propagation effect. Extra processing such as balloon type classification and text recognition are also covered. The third approach introduces a knowledge-driven system that combines low and high level processing to build a scalable system for comics image understanding. This approach is intended to improve the overall precision of content extraction methods. We built an expert system composed by an inference engine and two models, one for comics domain and another one for image processing, stored in an ontology. The first model embeds the knowledge about comic books and the second models the image processing related part. These two models allow consistency analysis of extracted information and inference of the relationships between all the extracted elements such as the reading order, the type of text (e.g. spoken, onomatopoeic, illustrative) and the relations between speech balloons and speaking characters. The expert system combines the benefits of the two first approaches and enables high level semantic description such as the reading order, the semantic of the balloon shapes, the relations between the speech balloons and their speakers, and the interaction between the comic characters.

Apart from that, in this thesis we have provided the first public comic book image dataset and ground truth to the community\* [1] along with an overall experimental comparison of all the proposed methods and some of the state-of-the-art methods† [2, 3].

## References

- [1] C. Guérin, C. Rigaud, A. Mercier, and al., “ebdtheque: a representative database of comics”, *IEEE Proceedings of 12<sup>th</sup> International Conference on Document Analysis and Recognition*, Washington DC, pp. 1145-1149, 2013.
- [2] C. Rigaud, C. Guérin, D. Karatzas, J.-C. Burie, J.-M. Ogier, “Knowledge-driven understanding of images in comic books”, *International Journal on Document Analysis and Recognition (IJDAR)*, Springer Berlin Heidelberg, ISSN 1433-2833, pp. 1-23, 2015.
- [3] C. Rigaud, “Segmentation and indexation of complex objects in comic book images”, *University of La Rochelle (France) and Autonomous University of Barcelona (Spain)*, ISBN 9788449049620, 2014.

\*<http://ebdtheque.univ-lr.fr>

†<http://www.christophe-rigaud.com>